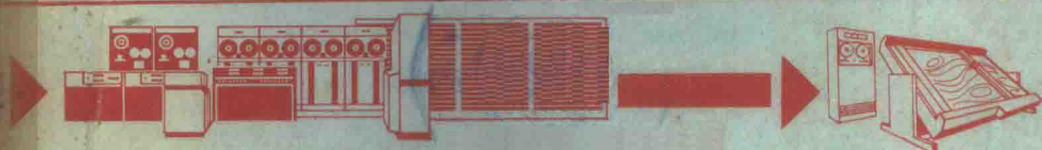


数学地质丛书

地质数据库

于志钧 编著



地质出版社

内 容 提 要

数据库是现代电子计算机系统应用的一个重要领域。近些年来,一些国家在工商企业管理、资料图书的检索、人民生活的社会服务等方面,都由于使用数据库而实现了完全的自动化,极大地提高了工作质量和效率。

本书以地质数据库为例介绍了初级数据库的基础知识、COBOL 语言、数据文件的程序设计与应用,并给出了典型程序及算例。最后一章还介绍了现代大型数据库的常识和发展现状。

本书适于广大地质人员和地质院校师生阅读,可据以设计初级地质数据库,也可供工农商企业、人民生活的社会服务单位、档案图书资料部门等设计和使用初级数据库的人员参考。

数学地质丛书

地 质 数 据 库

于志钧 编著

责任编辑 高书平

*

地质部书刊编辑室编辑

地质出版社出版

(北京西四)

地质印刷厂印刷

(北京安德路 47 号)

新华书店北京发行所发行·各地新华书店经售

*

开本: 850×1168 1/32 印张: 6¹¹/16 插页二个 字数: 174,000

1981 年 2 月北京第一版·1981 年 2 月北京第一次印刷

印数 1—3,110 册·定价 1.30 元

统一书号: 15038·新 635

前 言

正象由于电子计算机外部设备的日益发展、完善,电子计算机本身已经退居为现代电子计算机系统中的一个部件的地位一样,电子计算机仅仅作为计算工具这一功能,在现代电子计算机系统的庞大功能中也退居到一个不那么显赫的地位了。数据库(database)就是现代电子计算机系统应用的一个极其重要的发展领域。近十余年数据库的发展非常迅速,它使现代企业管理、商业经营、科学技术资料管理、各种经济技术情报的收集与检索、图书馆业务、人民生活的社会服务等等走向完全的自动化,大大提高了工作效率。数据库这门知识也发展成了一门新兴的独立学科。

本书就数据库在地质勘探方面的应用对初级数据库的基本知识、COBOL 程序设计语言、数据文件的程序设计和应用做了详细地介绍,文中例举了典型程序及计算机处理算例。

COBOL 是广泛应用于数据库设计的基础程序设计语言。在我国,由于过去使用计算机多限于科学计算,所以至今还没有专著介绍 COBOL 语言。本书从程序设计角度对 COBOL 语言作了详细的、系统的、也是最基本的介绍。今后随着我国电子计算机技术的迅速发展,COBOL 将日益发挥它在工农企业管理、商业经营、科技数据管理、科技情报检索、社会服务等方面的作用。

对于现代大型数据库,即包含数据通信、数据库网络的数据处理中心,本书仅介绍一些基本常识和国外发展现状。这方面的深入知识请参阅专门著作。

本书是一本深入浅出的数据库基础知识读物,文中不涉及高深的数学知识和电子学原理,具有中等以上文化水平,对计算机程序设计有初步了解的同志都可以看懂它。

本书适于地质科技生产和研究人员、大学地质专业师生及从

事数学地质工作人员阅读，也可供其它专业及经济管理人员参考之用。

本书承蒙中国科学院地质研究所刘承祚同志审阅并提出许多宝贵意见，对此致以衷心的感谢。

编著者

目 录

第一章 绪论	1
第一节 什么是地质数据库?	1
第二节 数据处理系统的硬设备.....	3
第三节 信息管理系统的基本概念.....	15
第二章 数据库程序设计基础语言—COBOL	21
第一节 COBOL 概念.....	21
第二节 标识部.....	25
第三节 环境部.....	26
第四节 数据部.....	27
第五节 过程部及 COBOL 指令.....	43
第三章 磁带文件	63
第一节 磁带记录.....	63
第二节 磁带文件组织.....	65
第三节 磁带文件的 COBOL 指令.....	66
第四节 磁带顺序文件的典型 COBOL 程序及其在地质数据 处理中的应用实例.....	71
第四章 磁盘文件	82
第一节 直接存取装置.....	83
第二节 磁盘记录.....	83
第三节 磁盘文件组织.....	85
第五章 直接存取文件	93
第一节 索引顺序文件的 COBOL 指令.....	93
第二节 索引顺序文件的典型 COBOL 程序及其在地质数据 处理中的应用实例.....	100
第三节 随机文件的 COBOL 指令.....	113
第四节 随机文件的典型 COBOL 程序及其在地质数据 处理中的应用实例.....	116

第六章 分类与汇总	130
第一节 分类与汇总的概念	130
第二节 分类方法	130
第三节 分类文件的 COBOL 指令	132
第四节 分类文件的典型 COBOL 程序及其在地质数据 处理中的应用实例	135
第七章 查表技术	151
第一节 查表方法	151
第二节 查表的 COBOL 指令	155
第三节 查表的 COBOL 程序及其在地质数据库中的应用实例	157
第八章 地质数据库的设计和应用	166
第一节 地质数据库的设备配置	166
第二节 地质数据库的设计方法	168
第三节 地质数据库的使用和管理	172
第四节 矿产预测数据库应用算例	173
第九章 现代数据库简介	186
第一节 现代数据库结构	187
第二节 现代数据库的存储结构	191
第三节 数据子语言	196
第四节 数据库网络	200
结 语	201
附录 COBOL 保留字	202

第一章 绪 论

第一节 什么是地质数据库？

地质数据库是地质资料(包括文字的和数字的以及图表资料)的电子计算机外存(磁带、磁盘)档案及检索系统,也就是实现地质资料的存储、更新和检索自动化。

地质资料的传统处理方式是,按照规定的格式填写成书面文件、报表并附图,然后装进档案袋,按地区、类别编号,放入资料柜中的编号抽屉内,保存备用。为了便于查找,另设一套目录卡片,在卡片上写明资料名称、编号、作者、存放地点的编号(资料柜号、抽屉号、档案袋号、文件号)。当使用者需要某项地质资料时,就根据目录卡片查出有关文件的存放地址,按地址编号找出文件,查阅所需资料,摘录后把文件复原。这就是一般资料室的概貌。

地质资料的电子计算机存储、更新和检索,在国外是从六十年代初开始的。它是随着数据管理和计算机地质软件两个方面的发展需要应运而生的。一个地质勘探部门从事从野外地质调查到室内化验、鉴定,从地面到井下等各方面工作,取得大量地质资料。为了对这些资料进行科学管理,建立了相应的资料室或档案馆。但随着科学技术的发展和地质勘探规模的扩大,地质资料的数量以极大的速度增长着,资料的管理、使用和更新便成了一项十分繁杂的工作,传统的简单的手工业式的资料管理方法已不能适应形势发展的需要;实现地质资料的电子计算机存储和管理完全是一件垂手可得的现成技术而其好处是显而易见的。另一方面,随着计算机技术广泛应用于地质学领域,计算机运算的高速度(每秒运算几十万、几百万、几千万次,甚至上亿次)和原始数据准备的低速度

的矛盾便出现了：用人工查找所需数据并穿孔卡片，然后输入计算机进行运算，显然与计算机本身(主机)和输出设备日益完善，一台计算机可以同时驱动几十台各种外部设备，同时为几十个用户服务这一高效能是极不相称的。为了改善资料管理和适应电子计算技术已广泛引入地质科学领域的需要，人们开始考虑建立统一的地质资料数据库。数据库这一新的科学，由于电子计算机外部存储装置(磁带、磁盘)的完善和电子计算机的文件管理功能日益增强，首先在商业部门发展起来的。1961年，美国几个石油公司首次把实际钻井的地质资料转换成机器存储形式，存储在磁带或磁盘上。

有人认为建立地质数据库花费人力物力太大，成本高，有些存入机器外存的地质资料很少有机会使用或根本用不着，使大量机器外存被无用的信息占据。后来的实践证明这些问题是不存在的。数据库一经建立，通过精心设计，数据库组织得非常严密，所存地质资料的利用率是高的。

地质数据库目前已经受到各技术先进国家的普遍重视，世界上已建立了450到500个各类地质数据库，涉及内容非常广泛，有综合性地质数据库，也有专用地质数据库。

在综合性地质数据库方面，英国设计了名为G-EXEC的综合数据处理系统，能够处理任何类型的地质数据，在一般的中型计算机上，可以处理专门地质数据，同时也能处理基础档案地质数据。这一系统的软件是用FORTRAN IV语言设计的，便于非程序地质人员掌握。

地质数据库起源于美国，美国应用地质数据库也最为普遍，其中著名的有：PDS(石油数据系统)，储存了美国和加拿大的石油与天然气田数据；CRIB(计算机资源信息库)，存储有关世界金属与非金属矿产资源数据；NCRDS(国家煤炭资源数据系统)；WHCS(井史控制系统)，包含80万口钻井记录，用COBOL语言设计的专用程序处理。

在加拿大、瑞典、芬兰、西德建立了野外地质数据库。法国发

展了名为 SIGMI 的地质数据存储、检索和处理系统。

此外,许多国家还在地球化学、地热资源、环境地质、水文地质等方面建立了专用地质数据库。

最近国外出现了利用地质数据库预报矿产资源的趋势,在矿床数据库的基础上建立矿床模型,用以预测所研究面积的含矿性。在这方面,土耳其(由美国科学家进行的)做了铬铁矿床、铁矾土镍矿床及铀矿资源的估计。美国自己也建立了一个叫作 PROSPECTOR(探矿者)的矿产勘探评价系统。这个系统目前存储了硫酸盐类型矿床、碳酸盐铅锌矿床及铜矿床的数据,计划最终存储 50 种不同矿产的数据,以预测矿产资源。

近年来,在我国随着电子计算机技术的发展与普及,电子计算机技术已被引入地质学领域,地质数据库已在一些有条件的地质科研、生产部门做了小规模试验,效果令人十分满意。可以预料,今后随着我国电子计算机外部设备制造工业的发展,数据库必将普遍建立起来,数据库这门科学也必将迅速地发展,为在我国实现社会主义现代化做出重要贡献。

第二节 数据处理系统的硬设备

数据库除用电子计算机存取和检索数据外,还包括数据加工,如数学运算、绘图、制表等等。所以说,数据库是一个数据处理系统,是以电子计算机(主机)为中心的现代电子计算机系统。

所谓硬设备就是电子—机械设。现代电子计算机系统由中央处理机、主存储器、控制台、基本输入—输出设备、显示装置、数据传送系统及终端装置组成(图 1-1)。

1. 主存储器

主存储器是计算机的脑细胞,由记忆元件组成,目前主要是由磁芯组成,所以也叫作磁芯体。它的功能是存储信息,也就是信号,由磁化和去磁两种物理状态来记忆 0 和 1 两种信息。这样的信息称作二进位制的位(bit)。主存储器由许多存储单元构成,每一单

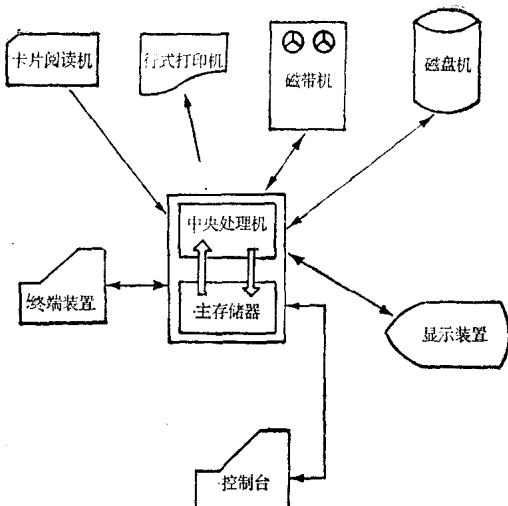


图 1-1 现代电子计算机系统连络示意图

元包含若干个二进位制位，计算机在运行时从主存取送的数据都是由二进位制位组成的。取送数据的术语叫作信息交换。现代电子计算机以八个二进位制位组成的字节(byte)作为信息交换的最小单位。二个、四个或八个字节组成一个计算机字(word)，常称为机器字。这样，一个字就包含 16、32 或 64 位。更大型计算机的字长有 128 位的。通常一个字就是一个存储单元，字长愈大，存储单元存放的数的位数就愈多，精度就愈高。主存的存储单元愈多，计算机的记忆能力就愈强，能够存储的数据就愈多。主存储器因为在计算机的内部直接与中央处理机交换信息，所以也叫作内存器，简称**内存**。主存储器含有的存储单元或字节数称作内存容量或计算机容量。主存储器的每个字节都有自己的编码，称作**地址**。所有指挥机器操作的**指令**和参加运算的数据都存放在主存中的存储单元之内，这就实现了计算机的记忆功能，并能有条不紊地按照计算程序的要求准确地存取数据和指令。

2. 中央处理机

中央处理机控制着整个计算机系统的操作。所有对主存的访

问①、运算和输入—输出操作都由中央处理机启动和执行。中央处理机受用户程序指挥。中央处理机只能执行存放在主存储器内的指令,不直接与外部存储器(磁带、磁盘)打交道,所以指令和数据在被中央处理机执行之前必须加载到内存,也就是送入主存储器。

3. 控制台

控制台是操作员与计算机的对话装置,由控制台打字机(简称控打机)控制。操作员可以通过它干预机器运行(如输入程序,中途改变程序的执行顺序,中断程序运行等),或机器要求人工干预(如程序出错,要求分配外部设备等)。机器运行状态和要求打印在记录纸上,操作员随时观察记录并通过按键给计算机下达命令。

图 1-2 是一台大型控制台打字机。

4. 基本输入——输出设备

基本输入—输出设备是指电子计算机最低限度要配备的输入—输出设备,通常有卡片阅读机(简称卡读机)、行式打印机、磁带机与磁盘机。这些设备有的单纯是输入用的,如卡读机;有的单纯是输出用的,如行式打印机;有的既是输入装置又是输出装置,如磁带机和磁盘机,所以我们把这类设备总称之为输入—输出设备。此外,还有一些输入辅助设备,如卡片穿孔机,它是脱离计算机的,但却是必不可少的。卡片穿孔机把大量的书写数据转换成计算机可以接受的卡片上的穿孔信息。

(1) 卡片阅读机 是光电输入机。输入时,使穿孔卡片通过光源,光线透过卡片上的孔射到光敏元件上,产生电流信号,经过放大电路产生电脉冲,表示“1”;在卡片没有孔的位置,光线被遮挡,不产生电信号,表示“0”。光电输入机就这样把穿孔卡片按列组成的二进位制编码信息逐列输入计算机内存。

现在穿孔卡片的规格已经标准化,国际上不同厂家生产的不同型号计算机都采用同一规格的卡片。

① 访问是计算机专用术语,就是对主存储器的存储单元按地址查找,好像访问(打听)。

标准卡片格式为 80 列卡片(图1-3)。每列有 0 到 9 及上部二行共 12 个穿孔位置。上部三行为区间孔, 其中一行占 0 的位置。无区间孔是数 0 到 9, 有区间孔是文字 A 到 Z 或特殊符号(如 +、-、*、/、=、>、<、()、:、,、'、%、&、\$ 等)。

图 1-4 是一台穿孔和符号两用卡读机, 读卡速度每分钟 90 张。

(2) 行式打印机 是电子计算机的基本输出装置。打印纸宽度标准规格为 15 英寸, 每行打印 132 个字符位。可以打空位, 即留空格。图 1-5 所示型号的行式打印机速度是每分钟打印 190 到 330 行, 字符种类有 64—128 种。打印纸走纸靠 8 单位穿孔纸带控制。

(3) 磁带机 是既可以存(写入)又可以取(读出)信息的装置, 是一种输入—输出设备。

磁带机的存储介质是磁带。磁带由塑料带基涂氧化铁层制成。当存放信息时, 铁氧层被磁化。磁带具有使用轻便, 成本低, 记录速度快, 容量(存储信息的数量)大, 可重复使用(当记录不需保存时可抹掉或直接记录新信息而把老信息冲掉)等优点。

磁带记录有 7 个磁道和 9 个磁道两种, 信息记录在磁道上, 由磁带机的读写磁头读或写。图 1-6 是磁带机。

(4) 磁盘机 是一种快速直接存取数据的装置, 是数据库的关键设备。磁盘是装在一个轴上的若干个平行圆盘(铝盘), 元盘表面涂有磁层, 它的上下两面都可以存储信息。盘面上分布有同心圆磁道, 二进位制信息记录在磁道上。工作时, 磁盘高速(2400 转/分)旋转, 取送数据靠存取臂内外相对移动, 由装在臂端的读写磁头在盘面上读或写信息。磁盘直径一般为 15 英寸, 每套磁盘组有 1 到 80 片磁盘, 可存储 2 至 800 百万字节。磁盘存取速度极快, 存取时间一般为 20 到 75 微秒(秒/百万), 所以我们使用磁盘时, 常常不感觉它是外部存储器, 似乎是扩大了主存的容量。

图 1-7 是一个 3 片磁盘机。它有 4 个读写磁头, 每个磁道分 12 段(扇形区), 每道可存储 1632 个数据记录, 总容量为 500 万字

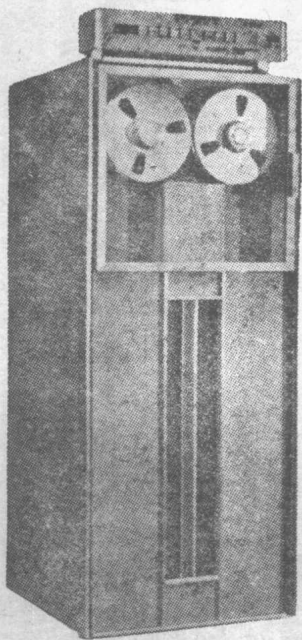


图 1-6 磁带机

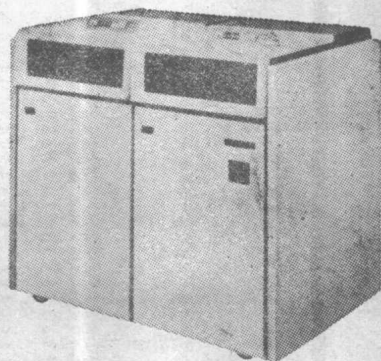


图 1-8 大型磁盘机

(word)。

图 1-8 是二台大型磁盘机。每台有 11 片磁盘，容量为 70 百万字节，存取速度为 20 微秒。

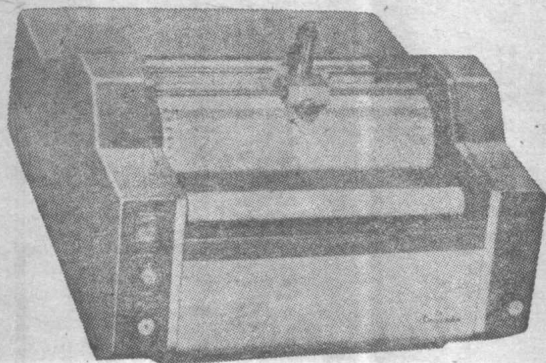


图 1-9 卷动式 X—Y 数控绘图机

5. 绘图设备

常用的绘图设备是X—Y 数控绘图机。它是一种电子机械绘图装置,有卷动式及平板式两种。

卷动式绘图机是将绘图纸卷在两个卷纸轴上,卷轴可以正反向转动,带动纸面上下移动,绘图笔在垂直纸面运动方向的滑架上往复移动。笔有抬笔(空走)及落笔(绘图)动作。通过纸面与绘图笔的垂向相对运动可以绘出任意形状的图件。这种绘图机的绘图幅宽为50—80厘米,绘图长度以轴上的卷纸量为限,所以实际上长度是不受限制的。图1-9是一台小型卷动式X—Y 数控绘图机。

平板式绘图机的结构如图1-10所示。绘图纸卷在绘图板两端的卷纸轴上,绘图板两边各有一条精密导轨,游梁架于导轨上,绘图笔架放在游梁架上,笔架装一至四支绘图笔,分红、兰、棕、黑各色。横游梁可沿导轨往复移动,笔架在横游梁上可垂直导轨方向移动,这样,绘图笔就可以到达纸面的任何位置,实现绘图。绘完图板上整个纸面后,用卷纸轴卷动纸面,将画过的纸面卷上,再绘新图。这种绘图机规格有小型的,图板面积约 30×50 厘米²,中型的约 80×100 厘米²,大型的达 3×5 米²。图1-10是一台大型平板数控绘图机及其绘制的等值线图。

X—Y 数控绘图机可直接由计算机驱动,即所谓联机绘图;也可由计算机把计算结果和绘图指令记在磁带上,然后把磁带取下装在专用脱机绘图机上,即脱机绘图。后者可以节省占用计算机的时间。在图1-10给出的是脱机绘图流程。

6. 显示装置

显示装置的主体是阴极射线管,它的英文缩写是CRT (Cathode Ray Tube),所以简称CRT。它类似电视机的显像管,可以在管的屏面上显示字符或图形。CRT 由电子计算机控制,电子枪发出电子束,电子束通过型板,集中在屏面,形成图像(图1-11)。

CRT 装置可作数据输入或信息输出用,或者同时起两种

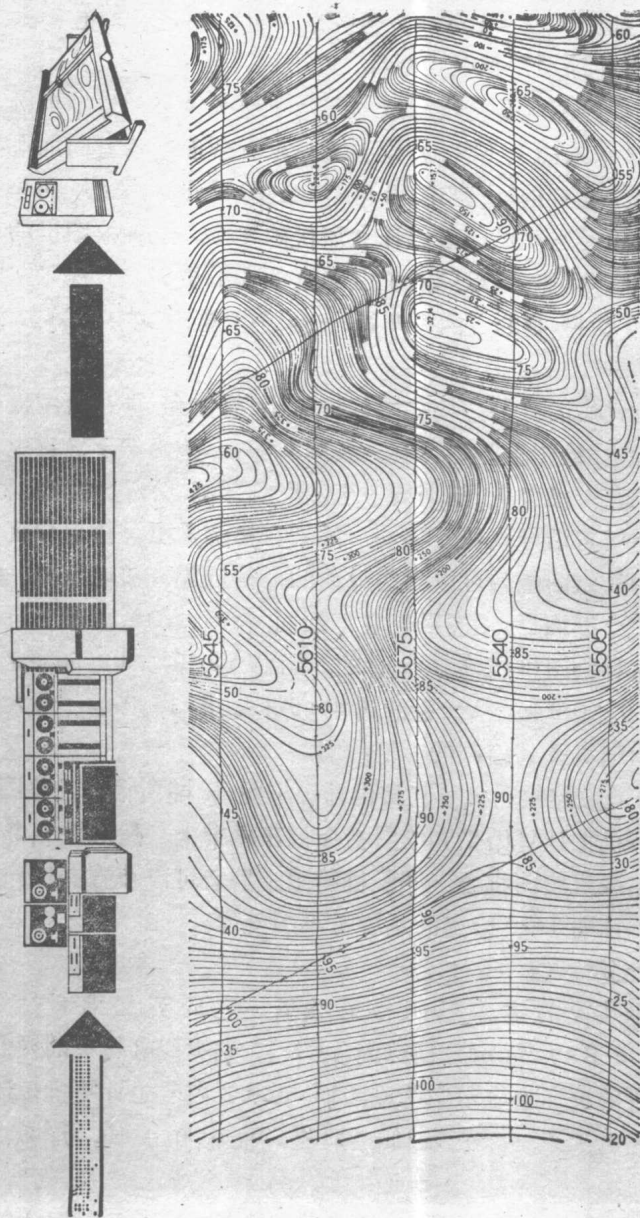


图 1-10 平板式 X—Y 数控绘图机绘图流程

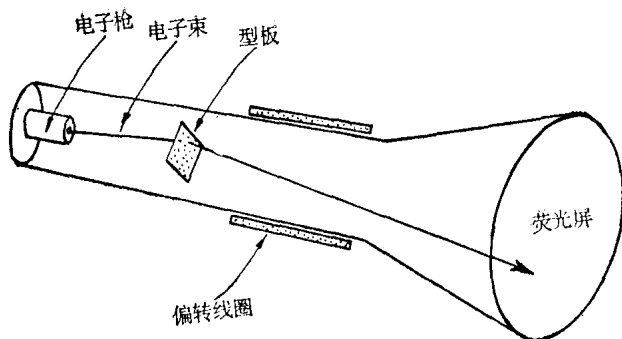


图 1-11 阴极射线管示意图

作用。

就其功能来说, CRT 装置可归为三大类:

(1) 信息检索 CRT 装置可以迅速地显示计算机运行状态、中间结果和计算结果。操作员随时可以通过它检索需要了解的过程信息。

(2) 数据输入 操作员可通过 CRT 装置的字符按键把数据信息送入显示装置的缓冲器(一种暂存装置), 并使其在屏幕上显示出来, 经检查无误后再按输入键送入计算机系统。这样输入的优点: 一是避免出错; 二是节省卡片。但因为它是联机设备, 不能同时为多个用户服务, 所以对大批用户来讲, 仍以穿孔卡片输入为主。

图 1-12 是终端字符显示装置, 屏面为 14 英寸, 可显示两种尺寸的字符: 1024 个字符(16 行, 每行 64 个字符); 2048 个字符(32 行, 每行 64 个字符), 字符种类 128 种。

(3) 交互绘图 是现代化工程设计广泛使用的一种自动绘图辅助设备。交互绘图设计是人对计算机系统干涉的对话操作, 使用叫作光笔的装置, 直接干涉计算机系统。

图 1-13 表示光笔从 CRT 屏面接收光脉冲。操作员使用控制板上的按键、开关, 指示交互绘图装置执行“绘图”、“抹掉”、“移动”等操作。光笔指到那里, 就等于告诉计算机所绘的这一部分图是要抹掉还是移动。绘图是由随着光笔路线的专门轨迹程序完成的。